

AD-A152 588

MUTUAL RECOGNITION OF HUMAN FALLIBILITY: A RESOLUTION
OF PRISONERS' DILEMMA(U) RAND CORP SANTA MONICA CA
J A CAVE JUL 84 RAND/P-7030

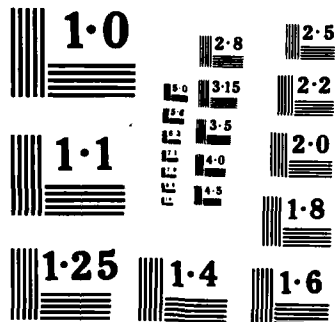
1/1

UNCLASSIFIED

F/G 12/1

NL

										END			
										FILED			
										DTIC			



AD-A152 588

2

MUTUAL RECOGNITION OF HUMAN FALLIBILITY:
. A RESOLUTION OF PRISONERS' DILEMMA

Jonathan A. K. Cave

July 1984

DTIC FILE COPY

DTIC
ELECTE
APR 17 1985
S D

E

This document has been approved
for release and sale in
unlimited.

85

P-7030

4

15

06.5

Introduction

In this note, we shall argue that mutual appreciation of human fallibility can resolve the Prisoners' Dilemma that so starkly describes many conflict situations. To do so, we shall suggest a new interpretation of an old idea: successive elimination of dominated strategies. We shall compare it to other solutions which reflect individual recognition of human fallibility, such as perfect and sequential equilibrium. Then we shall apply it to a repeated Prisoners' Dilemma in which players are allowed to base their actions only on the outcome of the previous round of play. While the results depend on whether players are allowed to react to their own previous moves or not, we shall find that the stubbornly noncooperative behavior that characterizes the one shot game is ruled out by the solution we propose. Most "resolutions" of Prisoners' Dilemma either include such noncooperative behavior as one of the possible outcomes, or are able to ensure cooperation only by the use of *ad hoc* beliefs or adjustment mechanisms specifically designed to favor cooperation. The "wide equilibrium" concept we shall propose is largely free of arbitrary specifications.

There are two fundamental ideas behind our proposal. The first is the individual recognition of human fallibility, and the second is that this fallibility will not only be recognized, but will become common knowledge. After brief discussions of these ideas, we give formal definitions and discuss the relation between the concept of trembling hand perfect equilibrium, which is based on individual recognition of human fallibility, and wide equilibrium, or equilibrium in sequentially undominated strategies. The final section applies this solution to the Prisoners' Dilemma.



Accession For	
NTIS GSA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

A. Individual Recognition of Human Fallibility

In noncooperative games, the issues of credibility and fallibility are closely connected. In a game in tree or extensive form, players are making threats or promises when they declare their strategies. A particular n -tuple of strategies is in equilibrium if no player can unilaterally change his strategy in such a way as to improve his payoff. However, a player might not believe that a change of strategy would be unilateral, if an opponent has made a threat that he would not wish to carry out. The requirement that all threats be credible in the sense that each player would actually wish to carry out all his threats if push came to shove is captured in the idea of subgame perfectness [Selten, 1973]. It reflects mutual appreciation of human rationality, since a threat is credible if a rational opponent would carry it out.

The issue of credibility does not naturally arise in games represented in normal or strategic form, where all players make a single simultaneous move. On the other hand, we can obtain a stronger perfectness concept by changing our focus from credibility to fallibility. Suppose that player A does not believe that player B will carry out a threat because it is not in B's interest to do so. If the players are in equilibrium, this means that player B is not called upon to carry out his threat unless player A changes his move. In the normal form of the game, player B commits himself to all the contingencies in his strategy when he declares it. Therefore, player A will know that the threat would be carried out automatically. In committing himself to the threat, player B is relying on player A's ability to avoid invoking it. If player B thought that player A would be unable to carry out his declared strategy (which avoids the threat) with perfect accuracy, he would not wish to commit himself to the costly threat.

While subgame perfectness means that players view each other's strategies as products of a rational mind, the idea of human fallibility outlined above suggests that no player should place too much reliance on the rationality of his opponents. The emphasis has shifted from acts to beliefs. In subgame perfectness, a player's acts must be rational given

the belief that the other players will follow their declared strategies for the "rest of the game." By contrast, the idea of fallibility rests on beliefs which are not derived from declared strategies.

There are several ways of capturing this idea in a solution. Perhaps the simplest is the concept of trembling hand perfectness [Selten, 1975]. An equilibrium is trembling hand perfect iff it is the limit of a sequence of completely-mixed n -tuples of strategies in which players are required to use inferior strategies with asymptotically negligible probability. The beliefs of the players are embedded in the relative probabilities with which strategies are used. Each player has the same model of independent mistakes or trembles by the other players, and the model is selected to fit a given equilibrium.

The idea that players might face situations in which their beliefs about prior or concurrent actions determine their choices can also be expressed in extensive-form games. If a player reaches a position he should not have reached had players been adhering to their declared strategies, he must have some beliefs as to what occurred. If he knows exactly what occurred, we may proceed as in the definition of subgame perfectness. If not, each player will need a model of what went wrong. The concept of sequential equilibrium [Kreps and Wilson, 1982a] uses an explicit set of beliefs to support behavior at "unreachable" positions. This is a weaker concept than trembling hand perfectness, since the beliefs are not required to entail completely mixed independent trembles by the other players. In addition, players are not required to use the same model of beliefs. Therefore, every trembling hand perfect equilibrium is subgame perfect, and every subgame perfect equilibrium is sequential.

Both sequentiality and trembling hand perfection support particular outcomes by particular beliefs. In addition, both sequentiality and subgame perfectness require players to evaluate the conduct of other players in the rest of the game according to their declared strategies, even in the face of evidence that other players are not adhering to

their declared strategies. Since everyone must know the beliefs of everyone else in order to satisfy the intuitive concept of credibility, it seems appropriate to look for a solution that invokes human fallibility without specifying the beliefs that players must hold.

If a player expects other players to use completely mixed strategies, he will never wish to put any weight on a weakly dominated strategy. This will be true no matter what mistakes he expects them to make. For this reason, no trembling hand perfect equilibrium uses such dominated strategies. We cannot go further without specifying the model of mistakes, so we shall define individual recognition of human fallibility by the requirement that *no player will use a dominated strategy*.

B. Mutual Recognition of Human Fallibility

The second idea upon which our proposal rests is that there are important differences between what each person knows and what is common knowledge [Aumann, 1976]. An event is common knowledge iff every finite statement of the form "A knows that B knows that . . . n knows whether the event has occurred" is true. One can regard common knowledge as the end result of a process of inference [Cave, 1983]. This is illustrated by the following story:

In a room, n people are gathered, $r \geq 2$ of whom are wearing red hats, while the others have blue hats. The hats were donned in the dark, so that no one knows the color of his own hat. After the lights are turned on, a bell is rung at one-minute intervals, and persons wearing red hats are instructed to leave the room at the next ring after they discover their hat's color. The bell rings for a long time, but nobody leaves the room.

Then a loudspeaker announces "There is at least one red hat in the room." As $n \geq 2$, this comes as no news to anyone. Nonetheless, r rings later, all the people wearing red hats leave the room. As a corollary, if by some chance all the people with red hats failed to leave the room, at the next ring all those wearing blue hats would leave.

What is important in this story is the fact that, while everyone knew that there were red hats in the room, it was not common knowledge until it was announced. Suppose $r = 2$. In this case, each person wearing a red hat would have observed the presence of one red hat. If person A sees that person B is wearing a red hat, he knows that either his own hat is blue, in which case person B should see no red hats, or person B sees a red hat, which must be person A's. Therefore, person A knows that if person B does not leave the room immediately after the announcement, his own hat must be red. The missing element was that A did not know that B knew that there was a red hat.

In noncooperative games, it is usual to assume that the strategy spaces and payoff functions are common knowledge. While this assumption is not necessary for Nash equilibrium, it does play a role in the definition of sequential equilibrium and perfectness concepts. Rather than requiring that all players share or recognize each other's beliefs, we propose the addition of an additional piece of common knowledge to the game: that no player will use a dominated strategy.

This announcement will trigger a chain of inference. Each person will already have discarded his or her dominated strategies. After hearing the announcement, they will no longer expect other players to use their dominated strategies. This means that the "rows" and "columns" corresponding to those strategies are deleted from the description of the game. In the reduced game, there may well be more dominations, and the process continues until no more can be found.

This idea is far from original with this paper. It appears, for example, in Luce and Raiffa [1959] as "strategies undominated in the wide sense". Farquharson [1969] terms it "sophisticated behavior under complete information," while Moulin [1979] considers "dominance solvable" games where iterated dominance reduces to a single outcome. However, this interpretation of the solution is new.

In the event that some players have more than one strategy undominated in the wide sense, we shall impose Nash equilibrium on the reduced game. This is the solution we shall call "wide equilibrium."

C. Perfection, Dominance, and Wide Equilibrium

We begin by defining the games and solution concepts we shall use.

Definition 1: a *finite game* is $[N, \Sigma, h]$, where

$N = \{1, \dots, n\}$ is the finite set of *players*;

$\Sigma = \Sigma_1 \times \dots \times \Sigma_n$ is the space of *pure strategy n-tuples*,
a Cartesian product of finite sets; and

$h: \Sigma \rightarrow \mathbb{R}^n$ is the *payoff function*

The space of *mixed strategies* of player i is a simplex of dimension $\#\Sigma_i$, denoted M_i . The product of the M_i is denoted M^N ; it is a subset of C , the simplex of dimension $\#\Sigma$, called the space of *correlated strategies*. If $\xi \in C$, then we denote by $\xi(s)$ the probability with which $s \in \Sigma$ is selected when ξ is played.

The space of correlated strategies of all players except player i is denoted C_{-i} . It is a simplex of dimension $\#\Sigma_{-i}$ where:

$$\Sigma_{-i} \equiv \prod_{j \neq i} \Sigma_j$$

We shall always use the subscript "-i" to denote the exclusion of player i . Thus an n -tuple s of pure strategies in which the strategy s_i of player i is replaced by $t_i \in S_i$ is denoted (t_i, s_{-i}) .

We can extend h to a payoff function $H: C \rightarrow \mathbb{R}^n$, by:

$$H(\xi) \equiv \sum_{s \in \Sigma} \xi(s) h(s)$$

Similarly, if $\mu = (\mu_1, \dots, \mu_n) \in M^N$, then $\mu(s_i)$ denotes the probability with which player i uses his pure strategy $s_i \in S_i$.

Finally, for each player we define an uppersemicontinuous *best reply* correspondence $\beta_i: C_{-i} \rightarrow M_i$, by:

$$\beta_i(\xi_{-i}) \equiv \arg \max \{H_i(\mu, \xi_{-i}): \mu \in M_i\}$$

When $\xi \in M_i \times C_{-i}$, there is no ambiguity in writing $\beta(\xi)$ for $\beta_i(\xi_{-i})$.

The product of the β_i is denoted β , and a *Nash equilibrium* is any fixed point of $\beta: M^N \rightarrow M^N$.

Definition 2: a Nash equilibrium m^* is *trembling-hand perfect* iff there exists a sequence μ^t of strictly positive members of M^N and a corresponding sequence ε^t of positive numbers with the following properties:

- i) $\varepsilon^t \rightarrow 0$ and $\mu^t \rightarrow m^*$ as $t \rightarrow \infty$;
- ii) for all t, i ; $\mu^t(s_i) \geq \varepsilon^t$ implies $s_i \in \beta_i(\mu^t)$

In words, a trembling-hand perfect (THP) equilibrium is a limit of " ε -perfect equilibria:" completely mixed strategy n -tuples in which every "inferior" pure strategy is used with probability $< \varepsilon$.

Definition 3: a mixed strategy $m_i \in S_i$ of player i is *undominated* iff for every $\mu_i' \in M_i$, $\mu_i' \neq m_i$, there exists $\xi_{-i} \in C_{-i}$ s.t.

$$H_i(m_i, \xi_{-i}) > H_i(\mu_i', \xi_{-i})$$

Definition 4: Let $U_i(N, M^N, H)$ denote the set of undominated strategies of player i in the game $[N, M^N, H]$, and let $U^N(N, M^N, H)$ be the Cartesian product of the $U_i(N, M^N, H)$. Since each Σ_i is a finite set, the sequence of games defined by:

$$\begin{aligned} \Gamma_0 &\equiv [N, M^N, H] \\ \Gamma_t &\equiv [N, U^N(\Gamma_{t-1}), H] \end{aligned}$$

has a unique limit attained in a finite number of iterations. This game is written $[N, M^W, H]$ where M^W is the space of mixed strategies *undominated in the wide sense*. The pure strategies in M^N are denoted Σ^W . An equilibrium of $[N, M^W, H]$ is a *wide equilibrium* of $[N, M^N, H]$.

In general, the conjectures players entertain about each other's moves should bear some relation to what they know about the game. The effect of the difference between common knowledge and what is merely known to everyone can be seen in the "Chain Store Paradox" analyzed by Selten [1978] and Kreps and Wilson [1982a]. In that example, a firm faces a sequence of potential rivals. If a rival enters, the firm will lose money in that market. It has available to it a Pareto inferior retaliatory strategy that can inflict losses on the entrant. If this game is played with a finite number of rivals (periods) under conditions of common knowledge, there is no possibility of deterring entry via the threat of retaliation. It is common knowledge that the last potential entrant cannot be deterred, since it cannot help the incumbent to retaliate. Therefore, the penultimate entrant will not be deterred, since the last rival will enter no matter what has happened in the past. This, too is common knowledge, and the serial game collapses to the one shot game.

On the other hand, deterrence is possible if the condition that payoffs are common knowledge is relaxed. This can be illustrated in a model with one incumbent, I, and two potential entrants, A and B. If B assigned a positive prior probability to a payoff function which rewards I for retaliation, then I could ensure that B would not enter the market by retaliating against A, thus causing B to revise his estimate of I's "toughness" upwards. Moreover, even if B did not hold this belief, I would be led to carry out the retaliation against A if he thought that B could be deterred. Going one step further, if A thought that I thought that B could be deterred, then A would anticipate that I would retaliate, and would be deterred. If the stakes are high enough, and if beliefs are thought to be sufficiently sensitive to observation, then deterrence is possible any time the payoffs are not common knowledge,

even if (as in the latter two cases above) they are known to everyone. The process of revision for this example is dealt with explicitly in Kreps and Wilson [1982a], and Milgrom and Roberts [1982].

If we start with the position that payoffs are common knowledge, there are still possibilities for inference as regards conjectured behavior by opponents. In particular, under conditions of less-than-perfect certainty about the ability or willingness of other players to carry out their declared strategies, adding as common knowledge the statement that "no player will use a dominated strategy" will initiate a chain of inference tending to a situation in which the strategies actually considered by a player will be strategies undominated in the wide sense. Inference and communication lead us to this solution without the necessity of requiring that all players share a common, independent model of each other's mistakes or even that we be able to write down an explicit set of beliefs that might support observed behavior.

Theorem 5: Let m^* be a trembling hand perfect equilibrium, and let s_i be a dominated strategy of player i . Then $m^*(s_i) = 0$.

Proof: If μ_i dominates s_i , and if $\mu_{-i} \in C_{-i}$ is strictly positive in each coordinate, $H_i(s_i, \mu_{-i}) < H_i(\mu_i, \mu_{-i})$. Therefore $\mu^n(s_i) < \epsilon^n$ for all n , and $\mu^*(s_i) = 0$. QED

Theorem 6: Let s^* be a pure strategy equilibrium of a two person game such that s_i^* is undominated. Then m^* is trembling-hand perfect.

Proof: Let m^ϵ be an ϵ -perfect equilibrium, and let us write:

$$m_i^\epsilon \equiv (1-\epsilon)s_i^* + \epsilon m_i'$$

for some completely mixed strategy m_i' . The payoff to player $j \neq i$ if he uses the strategy s_j can be written:

$$H_j(s_j, m_i^\epsilon) = (1-\epsilon)H_j(s_j, s_i^*) + \epsilon H_j(s_j, m_i')$$

The equilibrium strategy s_j^* is a best reply to s_i^* , and since it is undominated, there is some completely mixed strategy m_i' such that s_j^* is a best reply to m_i' as well. Therefore it is a best reply to m_i^ϵ . Finally, since player i is called upon to use his perturbed strategy m_i' with total probability ϵ , m^ϵ is an ϵ -perfect equilibrium. QED

In some circumstances, this characterization can be extended. For example, all completely mixed equilibria are trembling hand perfect, and none places positive probability on a weakly dominated strategy. However, it is easy to show by examples that i) not all equilibria in undominated strategies are trembling hand perfect; and ii) even pure strategy equilibria in undominated strategies may be imperfect if there are more than two players.

i) *Not all equilibria in undominated strategies are trembling hand perfect:*

	d	e	f
	+	+	+
a	1,2	3,0	0,0
	+	+	+
b	1,2	0,3	3,0
	+	+	+
c	1,0	2,0	2,3
	+	+	+

Example 1

There is an equilibrium in which player 1 uses undominated strategies "a" and "b" with equal probability, while player 2 uses "d" with probability 1. However, in any ϵ -perfect equilibrium where player 1 is indifferent between "a" and "b," neither can be a best reply.

ii) *Not all pure undominated strategy equilibria are perfect if $n > 2$:*

If a pure strategy is undominated there is some correlated strategy of the other players to which it is a best reply. However, that strategy may not be independent as required by ϵ -perfect equilibrium. In addition, the "perturbation" a player uses to make one player use a given undominated equilibrium strategy may differ from the "perturbation" required to make another player use his undominated equilibrium strategy. An example of this second type of failure is given by the three person game below.

	c	d
a *	1,1,1	0,0,-10
b	0,0,1	* 0,0,1

e

	c	d
a *	0,0,1	* 0,0,1
b *	0,0,1	0,0,-10

f

Example 2

The pure strategy equilibria are indicated with asterisks. Strategies "a" and "c" are weakly dominant for players 1 and 2 respectively, so the only possible pure strategy perfect equilibria are (a,c,e) and (a,c,f). Both "e" and "f" are undominated. However, (a,c,e) cannot be perfect. If player 1 plays "b" with probability ϵ (close to 0) and player 2 plays "d" with probability ω (close to 0), player 3's payoff if he plays "e" is $1 - 11(1-\epsilon)\omega$; his payoff if he plays "f" is $1 - 11\epsilon\omega$. He will therefore only choose "e" if $\epsilon \geq 1/2$.

Theorem 7: Every game $[N, M^N, H]$ possesses at least one wide equilibrium.

Proof: By construction, H is continuous on the compact and nonempty set M^W . Therefore, the best reply correspondence β^W for the game $[N, M^W, H]$ is uppersemicontinuous. In addition, the set of mixed strategies of player i dominated by a member of M_i^W is convex, so that β_i^W is convex-valued, and $[N, M^W, H]$ has an equilibrium m^* , which is also an equilibrium of the original game $[N, M^N, H]$. QED

Every wide equilibrium is *a fortiori* undominated, so every pure strategy wide equilibrium in a two-person game is perfect, although the Prisoners' Dilemma example of the next section shows that the converse is false. Also, not every wide equilibrium is perfect. In Example 1 above, all equilibria are wide equilibria, since no strategy is dominated. This includes the imperfect equilibrium in which player 1 uses "a" and "b" with probability 1/2 each, and player 2 uses "d" with probability one.

D. Prisoners' Dilemma with Finite Memory

In this section, we concentrate on an infinitely-repeated Prisoners' Dilemma game, in which players do not discount the future, and in which their strategies are constrained to depend only on the outcome of the previous round of play. We shall examine two variants of this example. In the first, a player's move at any stage depends only on his opponent's previous move. In the second, a player's move may depend on both his own and his opponent's previous moves. In each game we shall consider pure strategy Nash, perfect, and wide equilibria. This extends previous work by Aumann, Kurz, and the author [1978], in which it was shown that Tit-for-Tat is the unique wide equilibrium of the first variant.

The Prisoners' Dilemma we shall use is:

	L	R
T	1,1	4,0
B	0,4	3,3

It has a unique equilibrium in dominant strategies at (T,L). The Appendix extends the reactive memory analysis to a general Prisoners' Dilemma.

One Period ("Reactive") Memory of Opponent's Move

If this game is played as an undiscounted supergame in which a player's move is allowed to depend on his opponent's last move, we obtain a normal form game in which each player has eight pure strategies. A strategy for player 1 [2] is written (a,b,c) [(d,e,f)] where:

- a [d] is the move in the first period;
- b [e] is the move when the opponent's previous move was R [B]; and
- c [f] is the move when the opponent's previous move was L [T].

As written, each player has two redundant strategies. Since we are concerned with long-term average payoffs, the noncooperative strategies (B,T,T) [(R,L,L)] and (T,T,T) [(L,L,L)] are the same, as are the cooperative strategies (T,B,B) [(L,R,R)] and (B,B,B) [(R,R,R)]. The payoff matrix for the reduced game is shown below. It has two pure strategy equilibria: [(B,B,T),(R,R,L)] with payoff (3,3), and [(T,T,T),(L,L,L)] with payoff (1,1).

	L,L,L L,L,R L,R,L R,R,L R,R,R			
T,T,T	1, 1	4, 0	1, 1	4, 0
T,T,B	0, 4	2, 2	2, 2	4, 0
T,B,T	1, 1	2, 2	1, 1	2, 2
B,T,B	0, 4	0, 4	2, 2	2, 2
B,B,T	1, 1	2, 2	2, 2	2, 2
B,B,B	0, 4	0, 4	3, 3	3, 3

The payoff matrix for the reduced game

	L,L,L L,L,R R,R,L R,R,R			
T,T,T	1, 1	4, 0	1, 1	4, 0
T,T,B	0, 4	2, 2	2, 2	4, 0
B,B,T	1, 1	2, 2	3, 3	3, 3
B,B,B	0, 4	0, 4	3, 3	3, 3

Stage 2

Each player has two dominated strategies. The initially cooperative "Tit-for-Tat" strategy that responds to cooperation with greed and *vice versa*, (B,T,B) [(R,L,R)], is dominated by the initially greedy Tit-for-Tat strategy (T,T,B) [(L,L,R)]. Also, the initially greedy "Tit-for-Tat" strategy (T,B,T) [(L,R,L)] is dominated by the initially cooperative Tit-for-Tat strategy (B,B,T) [(R,R,L)]. As neither dominated strategy is an equilibrium strategy, both equilibria are trembling hand (and thus subgame) perfect. Eliminating dominated strategies, we get the Stage 2 matrix.

At Stage 2, the "appeasement" strategy (B,B,B) [(R,R,R)] is dominated by the Tit-for-Tat strategy (B,B,T) [(R,R,L)]. This leads to Stage 3 (below) at which Tit-for-Tat dominates the initially greedy Tit-for-Tat strategy (T,T,B) [(L,L,R)]. Finally, at Stage 4 Tit-for-Tat dominates the unrelentingly noncooperative strategy (T,T,T) [(L,L,L)]. Therefore, the only wide equilibrium of this game is the Pareto optimal Tit-for-Tat equilibrium [(B,B,T),(R,R,L)].

	L,L,L L,L,R R,R,L			
T,T,T	1, 1	4, 0	1, 1	
T,T,B	0, 4	2, 2	2, 2	
B,B,T	1, 1	2, 2	3, 3	

Stage 3

	L,L,L R,R,L	
T,T,T	1, 1	1, 1
B,B,T	1, 1	3, 3

Stage 4

One Period ("Reactive-Signalling") Memory of Both Players' Moves

In this version of the game each player has five information sets, and therefore 32 pure strategies. The generic pure strategy can be written (v,w,x,y,z) , where:

v is the player's initial move;
 w is the player's move if the previous moves were (T,R)
 x is the player's move if the previous moves were (T,L)
 y is the player's move if the previous moves were (B,R)
 z is the player's move if the previous moves were (B,L)

As in the reactive memory case, there are some redundant strategies: (B,B,B,B,B), (B,T,B,B,B), (B,B,T,B,B), and (B,T,T,B,B) are all strategies of perpetual cooperation, as are (R,R,R,R,R), (R,L,R,R,R), (R,R,R,L,R), and (R,L,R,L,R). Similarly, (T,T,T,T,T), (T,T,T,T,B), (T,T,T,B,T), (T,T,T,B,B); (L,L,L,L,L), (L,L,L,L,R), (L,L,R,L,L), and (L,L,R,L,R) are all perpetually noncooperative strategies.

Associated with any pair of strategies is a cycle of at most four outcomes that will result if the strategies are used. As this is an undiscounted supergame, the order of play within the cycles does not matter, and the cycles provide a convenient way of presenting the payoff matrix. The cycles and associated payoffs are shown below, together with the ranking for each player.

label	cycle	payoff	label	cycle	payoff	label	cycle	payoff
a	BR	3 , 3	f	BR,TR	7/2, 3/2	k	BR,BL,TR	7/3, 7/3
b	BL	0 , 4	g	BR,TL	2 , 2	l	BR,BL,TL	4/3, 8/3
c	TR	4 , 0	h	BL,TR	2 , 2	m	BR,TR,TL	8/3, 4/3
d	TL	1 , 1	i	BL,TL	1/2, 5/2	n	BL,TR,TL	5/3, 5/3
e	BR,BL	3/2, 7/2	j	TR,TL	5/2, 1/2	o	all	2 , 2

Player 1: $c > f > a > m > j > k > h = o = g > n > e > l > d > i > b$

Player 2: $b > e > a > l > i > k > h = o = g > n > f > m > d > j > c$

The following table shows the outcomes as a function of the nonredundant strategies, which have been numbered for convenience.

v	L R L R L R L R L R L R L R L R L R L R L R
w	R R R R R R R R R R R R R R R L L L L L L L L
x	R R R R L L L L R R R L L L L R R R L L L L R L L
y	R R R R R R R R R L L L L L L R R R R R R R L L L L
z	R R L L R R L L R L L R R L L R L L R R L L L R L L
<hr/>	
v,w,x,y,z	1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2
	1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6
<hr/>	
1	T,B,B,B,B a a e e a a e e a b b a b b b a e e a a e e b b b b
2	B,B,B,B,B a a e e a a e e b b b b a b b a e e a a e e b a b b
3	T,B,B,B,T f f g g k k g g f g g b b g b f 1 1 k k 1 1 b b b b
4	B,B,B,B,T f f g g k k g g b b g b b b g f 1 1 k k 1 1 b b b b
5	T,B,B,T,B a a k k a h h h a l l a a l l a k k h h h h i i i i
6	B,B,B,T,B a a k k h h h h a l l a a l l a k k h h h h i i i i
7	T,B,B,T,T f f g g h h h h f g g o o g g f o o h h h h i i i i
8	B,B,B,T,T f f g g h h h h f g g o o g g f o o h h h h i i i i
9	T,B,T,B,B a c e c a a e e a b c a a b b a e c a a e e c b b b
10	T,B,T,B,T c c g c m m g g c g c m m g g c l c o o l l c b b b
11	B,B,T,B,T c c g g m m g g b b g b m b g c l l o o l l b b b b
12	T,B,T,T,B a c c c a a o o a l c a a l l c c c n n n n c i i i
13	B,B,T,T,B c a c c a a o o a l l a a l l c c c n a n n i i i i
14	T,B,T,T,T c c g c m m g g c g c m m g g c c c n n n n c i i i
15	B,B,T,T,T c c c g m m g g c g g m m g g c c c n n n n i i i i
16	T,T,B,B,B a a e e a a e e a b b b b b b d d d d a d e b b d b
17	T,T,B,B,T f f m m k k o o f m m b b b b b d d d d k d d d b d b
18	B,T,B,B,T f f m m k k o o b b m b b b b f d d k k d d d b b d
19	T,T,B,T,B a a k k h h h h a o o n n n n d d k d h h h h d d d d
20	B,T,B,T,B a a k k h h h h a o o n a n n a k k h h h h d a d d
21	T,T,B,T,T f f m m h h h h f m m n n n n d d d d h h h h d d d d
22	B,T,B,T,T f f m m h h h h f m m n n n n f d d h h h h d d d d
23	B,T,T,B,T c c c c j j j j b b c b j b j c d d d d d d d d b d
24	B,T,T,T,B c a c c j j j j c c c j j j j c c c d a d d d a d d
25	T,T,T,T,T c c c c j j j j c c c j j j j d d c d d d c d d d
26	B,T,T,T,T c c c c j j j j c c c j j j j c c d d d d d d d d

Outcomes for the one-period reactive-signalling memory game

This game has 26 pure equilibria with the cooperative outcome "a;" 4 pure equilibria with the noncooperative outcome "d;" and 2 pure equilibria with the semicooperative outcome "h" (players take turns cooperating). None uses a weakly dominated strategy, so all are trembling hand perfect. They are:

Cooperative equilibria: $\{5,12,13\}^2$; $\{6,13,20,24\}^2$; (6,12); (12,6).

Noncooperative equilibria: $\{25,26\}^2$.

Semi-cooperative equilibria: (21,22); (22,21).

The following table shows the equivalence between strategies of the two players, the stages at which they are eliminated, and the strategies that dominate them.

Stage	Dominated strategy			Dominating strategy	
	Player 1	Player 2		Player 1	Player 2
1	(T,B,B,B,B)	1 (L,R,R,R,R)	=	(T,B,T,B,B)	9 (L,R,R,L,R)
1	(B,B,B,T,T)	8 (R,R,L,R,L)		(T,B,B,T,T)	7 (L,R,L,R,L)
1	(B,B,T,B,T)	11 (R,R,R,L,L)		(T,B,T,B,T)	10 (L,R,R,L,L)
1	(T,T,B,T,B)	19 (L,L,L,R,R)		(B,T,B,T,B)	20 (R,L,L,R,R)
1	(B,T,T,B,T)	23 (R,L,R,L,L)		(B,T,T,T,T)	26 (R,L,L,L,L)
2	(B,B,B,B,B)	2 (R,R,R,R,R)		(B,B,T,T,B)	13 (R,R,L,L,R)
3	(T,B,T,B,B)	9 (L,R,R,L,R)		(B,B,T,T,B)	13 (R,R,L,L,R)
3	(T,B,T,T,B)	12 (L,R,L,L,R)		(B,B,T,T,B)	13 (R,R,L,L,R)
3	(T,T,B,B,T)	17 (L,L,R,R,L)		(B,T,T,T,B)	24 (R,L,L,L,R)
3	(B,T,B,B,T)	18 (R,L,R,R,L)		(B,T,T,T,B)	24 (R,L,L,L,R)
3	(T,T,T,T,T)	25 (L,L,L,L,L)		(B,T,T,T,B)	24 (R,L,L,L,R)
3	(B,T,T,T,T)	26 (R,L,L,L,L)		(B,T,T,T,B)	24 (R,L,L,L,R)
4	(B,B,B,T,B)	6 (R,R,L,R,R)		(B,T,B,T,B)	20 (R,L,L,R,R)
4	(T,T,B,B,B)	16 (L,L,R,R,R)		(B,B,T,T,B)	13 (R,R,L,L,R)
5	(B,B,B,B,T)	4 (R,R,R,R,L)		(B,B,T,T,B)	13 (R,R,L,L,R)
6	(T,B,T,T,T)	14 (L,R,L,L,L)		(B,B,T,T,T)	15 (R,R,L,L,L)
6	(T,B,B,B,T)	3 (L,R,R,R,L)		mixed domination	

The matrix corresponding to strategies undominated in the wide sense is:

v		L	R	L	R	R	R	L	R	R
w		R	R	R	R	R	L	L	L	L
x		L	L	R	L	L	L	L	L	L
y		R	R	L	L	L	R	R	R	L
z		R	L	L	R	L	R	L	L	R
v,w,x,y,z		5	7	10	13	15	20	21	22	24
5	T,B,B,T,B	*a	h	l	*a	l	h	h	h	i
7	B,B,B,T,T	h	g	g	o	g	h	h	h	i
10	T,B,T,B,T	m	g	g	m	g	o	l	l	b
13	B,B,T,T,B	*a	o	l	*a	l	*a	n	n	*a
15	B,B,T,T,T	m	g	g	m	g	n	n	n	i
20	B,T,B,T,B	h	h	o	*a	n	*a	h	h	*a
21	T,T,B,T,T	h	h	m	n	n	h	d	*h	d
22	B,T,B,T,T	h	h	m	n	n	h	*h	d	d
24	B,T,T,T,B	j	j	c	*a	j	*a	d	d	*a

* = wide equilibrium

This game has 11 fully cooperative wide equilibria. In each of them, players will play Tit-for-Tat if their own previous move was cooperative. In addition, both of the semicooperative equilibria with outcome "h" are

wide equilibria in which players play Tit-for-Tat if their own previous move was noncooperative. There are no wide equilibria with the noncooperative outcome "d," the Pareto-inferior equilibrium of the one-shot game. Indeed, the noncooperative equilibria are eliminated earlier in this game than in the reactive memory game.

In any event, wide equilibrium resolves the Prisoners' Dilemma for the one-period reactive memory game. There are Pareto-inferior wide equilibria in the reactive-signalling memory game, but they are neither as inefficient nor as compelling as the inefficient dominant strategy equilibrium of the one shot game. It remains an open question to what extent this resolution by wide equilibrium depends on limited memory. In particular, it would be interesting to compute the wide equilibria of the full supergame for the reactive memory and reactive-signalling memory cases.

The learning interpretation of wide equilibrium makes this example an attractive metaphor for the evolution of cooperation, in which increasing common knowledge about which strategies will not be used leads to the elimination of the worst kind of noncooperative behavior.

Since the chain of inference leading to wide equilibrium was triggered by the common knowledge statement that no player would use a dominated strategy, and since the use of dominated strategies is ruled out by expectations of human fallibility, it is tempting to view the realistic cooperation involved in the wide equilibrium strategies as a result of mutual recognition of human fallibility.

Appendix: Reactive Memory in the General Case

In this section, we show that the Tit-for-Tat cooperative strategies form the only wide equilibrium in a general Prisoners' Dilemma game. The payoff function is:

	L	R
T	α, α	β, γ
B	γ, β	δ, δ

where

$$(1) \quad \beta > \delta > \alpha > \gamma, \text{ and}$$

$$(2) \quad 2\delta \geq \beta + \gamma$$

Assumption (1) guarantees that the unique dominant-strategy equilibrium of the one-shot game is Pareto-inferior, and assumption (2) ensures that the "cooperative" outcome BR is Pareto optimal.

Players are allowed to react to their opponent's previous moves. After redundant strategies are eliminated, each player has six pure strategies described in the same way as in Section D above. There are seven possible outcome cycles. These cycles, and the resulting payoffs are:

label	cycle	payoff	label	cycle	payoff
a	BR	(δ, δ)	g	BR, TL	$(\alpha + \delta, \alpha + \delta)/2$
b	BL	(γ, β)	h	BL, TR	$(\beta + \gamma, \beta + \gamma)/2$
c	TR	(β, γ)	o	all	$(\alpha + \beta + \gamma + \delta, \alpha + \beta + \gamma + \delta)/4$
d	TL	(α, α)			

The players rank these cycles as follows:

$$\text{player 1: } c > a > \max\{h, g\} \geq o \geq \min\{h, g\} > d > b,$$

$$\text{player 2: } b > a > \max\{h, g\} \geq o \geq \min\{h, g\} > d > c$$

The matrix for the game is shown below (* = pure strategy equilibrium).

	(L,L,L)	(L,L,R)	(L,R,L)	(R,L,R)	(R,R,L)	(R,R,R)
(T,T,T)	d*	c	d	c	d	c
(T,T,B)	b	g	o	c	o	c
(T,B,T)	d	o	d	o	h	a
(B,T,B)	b	b	o	g	o	c
(B,B,T)	d	o	h	o	a*	a
(B,B,B)	b	b	a	b	a	a

On the first round, (B,T,B)[(R,L,R)] is dominated by (T,T,B)[(L,L,R)],
 (T,B,T)[(L,R,L)] is dominated by (B,B,T)[(R,R,L)] if $\beta + \gamma \geq 2\alpha$, and
 (T,B,T)[(L,R,L)] is dominated by (T,T,T)[(L,L,L)] if $2\alpha \geq \beta + \gamma$.

Therefore, the matrix of undominated strategies is:

	(L,L,L)	(L,L,R)	(R,R,L)	(R,R,R)
(T,T,T)	d*	c	d	c
(T,T,B)	b	g	o	c
(B,B,T)	d	o	a*	a
(B,B,B)	b	b	a	a

Both pure strategy equilibria are trembling hand perfect by virtue of Theorem 6. In this game, the naive cooperative strategy (B,B,B) [(R,R,R)] is dominated by the more realistic Tit-for-Tat strategy, leaving

	(L,L,L)	(L,L,R)	(R,R,L)
(T,T,T)	d*	c	d
(T,T,B)	b	g	o
(B,B,T)	d	o	a*

Player 1[2] prefers c[b] to g, so (T,T,B)[(L,L,R)] is dominated by (T,T,T)[L,L,L] if the players prefer d to o, i.e. if

$$(3) \quad 4\alpha \geq \alpha + \beta + \gamma + \delta$$

By the same token, both players prefer a to o, so (T,T,B)[(L,L,R)] is dominated by (B,B,T)[(R,R,L)] if the players prefer o to g, i.e. if

$$(4) \quad \alpha + \beta + \gamma + \delta \geq 2(\alpha + \delta) > 4\alpha$$

Finally, suppose that

$$(5) \quad 2(\alpha + \delta) > \alpha + \beta + \gamma + \delta > 4\alpha$$

In this case, $(T,T,B)[(L,L,R)]$ is dominated by a mixed strategy, which can be written $\lambda(T,T,T) + (1-\lambda)(B,B,T)$ $[\lambda(L,L,L) + (1-\lambda)(R,R,L)]$ for some $\lambda \in (0,1)$. W.l.o.g., we shall concentrate on player 1's payoff. The mixed strategy clearly is a best reply to (L,L,L) , so there are two conditions for it to dominate (T,T,B) :

$$(6) \quad \alpha + \delta \leq \lambda\beta + (1-\lambda)(\alpha + \beta + \gamma + \delta)$$

$$(7) \quad \alpha + \beta + \gamma + \delta \leq \lambda\alpha + (1-\lambda)\delta$$

The condition for the existence of such a mixed strategy is therefore:

$$(8) \quad (\delta - \alpha)(\beta - (\alpha + \delta)/2) \geq (\chi - \alpha)(\beta - \chi), \text{ where}$$

$$\chi \equiv (\alpha + \beta + \gamma + \delta)/4$$

(5) implies that $(\delta - \alpha) > 2(\chi - \alpha)$, and (1) implies that $2\beta - \alpha - \delta > \beta - \chi$, so that (8) is automatically satisfied. Therefore $(T,T,B)[(L,L,R)]$ is dominated, and the strategy matrix reduces to:

	$(L,L,L) \mid (R,R,L)$	
(T,T,T)	d^*	d
(B,B,T)	d	a^*

in which (B,B,T) dominates (T,T,T) . Therefore, the only wide equilibrium is the cooperative Tit-for-Tat equilibrium $(B,B,T),(R,R,L)$.

REFERENCES

- Aumann, R.J.,: "Agreeing to Disagree," *Annals of Statistics*, 4, 1976, pp. 1236-9.
- Aumann, R.J., Kurz, M. and J. Cave,: "Prisoners' Dilemma with Memory 0," mimeo, 1978.
- Cave, J.,: "Learning to Agree," *Economics Letters*, 12, 1983, pp. 147-52.
- Farquharson, R.,: *Theory of Voting*, New Haven, Yale University Press, 1969.
- Kreps, D., and R. Wilson,: "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 1982a, pp. 253-79.
- Kreps, D., and R. Wilson,: "Sequential Equilibria," *Econometrica*, 50, 1982b, pp. 863-94.
- Luce, R. and H. Raiffa,: *Games and Decisions*, New York, John Wiley, 1959.
- Milgrom, P. and Roberts, J.,: "Predation, Reputation, and Entry Deterrence," *Journal of Economic Theory*, 27, 1982, pp. 280-312.
- Moulin, H.,: "Dominance Solvable Games," *Econometrica*, 47, 1979, pp. 1337-51.
- Myerson, R.,: "Refinements of the Nash Equilibrium Concept," *International Journal of Game Theory*, 7, 1978, pp. 73-80.
- Selten, R.,: "A Model of Oligopoly where Four are Few and Six is Many," *International Journal of Game Theory*, 2, 1973, pp. 141-201.
- Selten, R.,: "A Reexamination of the Perfectness Concept for Games in Extensive Form," *International Journal of Game Theory*, 4, 1975, pp. 2-55.
- Selten, R.,: "The Chain Store Paradox," *Theory and Decision*, 9, 1978, pp. 127-59.

END

FILMED

5-85

DTIC